

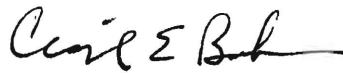
U.S. Economic Correlation
An Econometric Analysis of the Interdependency of the Global Economy

An Honors Thesis (Honors499)

by

Kevin Bergen

Dr. Cecil Bohanon

A handwritten signature in black ink, appearing to read "Cecil E. Bohanon". The signature is fluid and cursive, with a long horizontal stroke at the end.

**Ball State University
Muncie, Indiana**

May 2013

Expected Date of Graduation

May 2013

Sp2011
Undergrad
Thesis
LD
2489
.24
2013
1B47

Abstract

As the world becomes increasingly interconnected, the idea of domestic economies has faded and in its place the idea of a global economy has risen. Over the past ten years, the US has experienced a variety of economic ups and downs, including the recent 2008 recession. As the US economy has experienced such changes, the question remains as to how the rest of the global economy has also been affected? As such, a statistical analysis of the US and nine other countries (Brazil, China, India, Italy, Germany, Japan, Mexico, Peru, and the United Kingdom) based upon broad economic metrics will demonstrate the degree of interdependence between the US and the global economy. Basic econometric tools will be used to analyze the results of this research and I will discuss the most significant results from this process.

Acknowledgements

I would like to take the time here to thank my advisor, Dr. Cecil Bohanon, for guiding me throughout this entire process. He constantly helped me keep my ideas in order and gave a direction to all of my wild ideas. He has also been a fantastic professor and mentor for me since I have known him

I want to thank Dr. Lee Spector as well for all of his support over the past four years, always pushing me to try my hardest and achieve great things.

I would like to thank my family, Caron, Katie, Kyle, and Rob for listening to all of my ideas and economic discussions.

Finally, thanks to all my friends and professors here at Ball State University for your help and guidance.

Over the years, the economy has undergone periods of drastic shifts and swings. Some of these have been economic rallies; others have been drastic slumps, such as the great depression. The most recent example of a sustained economic downturn occurred in 2008, on the heels of the housing market crash and financial crisis. Economies all over the world struggled to grow and suffered long term periods of poor GDP growth. As the effects of the recession were felt almost everywhere worldwide, and the global economy becomes ever connected, the question arose as to how closely linked foreign economies are to the U.S. economy? The purpose of this project is to do an introductory econometric analysis of the correlation between select countries and the U.S. economy based on a variety of metrics.

This project, investigates ten countries, at different stages of economic development. Having this variety of countries would allow for a more comprehensive analysis. The data used is from sources such as IMF and WorldBank; using yearly data dating back to 2001, providing ten years of data. The metrics that were chosen to be part of the model are the more widely discussed metrics: GDP, GDP per capita, unemployment rate, and average income. These metrics provide a broad snapshot of the overall health of a country's economy and indicate whether or not it is growing or not. All of this data is in the appendix attached to this paper, as well as graphs for each individual metric. After data collection, E-Views 7 software was used to create simple linear regressions and provided econometric statistics of all the different data. I will detail the different regression equations that are used in for each model and discuss the significance and meaning of each of the variables and their estimates. In all cases, U.S. GDP was used as the dependent variable with different independent variables used based upon countries' economic metrics. For all regression models used in this paper, the initial discussion

will focus on the unrestricted model, using all the possible independent variables, and later removing the insignificant variables to yield the final restricted regression model.

One important thing that must be noted is the nature of this research as well as the results of the regressions. Throughout this project, I am testing for correlation, which shows how similarly two variables have trended over time. The regression results will provide us with this sort of result, showing how well independent variables can account for changes in U.S. GDP. As such, it must be said that correlation in no way means causality. Simply because a variable predicts changes in U.S. GDP does not mean that it is the reason for the change. For the most pertinent and appropriate regressions, I will attempt to explain why those variables are well suited to predict U.S. GDP, but again I can only prove correlation, not cause.

The 10 countries that were selected for this analysis are Brazil, China, India, Italy, Germany, Japan, Mexico, Peru, the United Kingdom, and the United States. When compiling regressions, they were focused on regional combinations that would tie U.S. GDP changes within a specific region or single country. While many regressions have been compiled, only the top 3 regressions, based upon their adjusted R-squared value, will be discussed in depth, the rest are available in Appendix A. In addition to these three regressions, there will be discussion of a regression on internal U.S. data, examining how well GDP growth rate, lag GDP, unemployment rate, and time predict U.S. GDP, as well as discussion of a regression that yielded a poor adjusted R-squared value.

U.S. Internal Regression Model

One simple regression to be considered is how well internal U.S. economic data is able to predict changes in U.S. GDP. In theory, the U.S. economy is highly dependent upon various internal factors such as unemployment rate and prior GDP levels. Thus, the below regression looks at internal variables to see how well they explain U.S. GDP:

USA Internal Regression Results, Table 1:

Variable	Unrestricted Model	Restricted Model
US Unemployment	3.15E+11 (1.55E+11)	N/A
Time	-7.18E+11 (3.53E+11)	N/A
GDP Growth Rate	2.99E+11* (7.90E+10)	1.39E+11* (3.23E+09)
US Lagged GDP	2.257* (0.5979)	1.043* (0.0054)
Constant	-1.47E+13 (6.93E+12)	-5.91E+11 (7.45E+10)
Adjusted R-Squared	0.9998	0.9997

n=10, numbers in parentheses are standard errors

** indicates significance at 5% level*

Data for group can be found in Appendix A

The data in Table 1 is for the years 2001-2011. The first model, or unrestricted model, looks at four variables, US unemployment, time, the US GDP growth rate, and US lagged GDP (lag is “t-1”). US unemployment is the percentage of workers currently unemployed, as defined by the US government. Time is a simple variable ranging from 1-10 based upon which year the data pertains to. The US GDP growth rate is a percentage value calculated by taking the change in GDP between two time periods, i.e. from “t-1” to “t”, GDP growth was 5.00%. US lagged GDP is used by comparing using the “t-1” value of GDP as a variable for GDP at time “t”. These four variables are then used in a regression equation as follows:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 USunemployment + \hat{\beta}_2 time + \hat{\beta}_3 rUSGDP + \hat{\beta}_4 lagUSGDP + u$$

In the above equation, β represents the estimated coefficients for each of the variables based upon the completed regressions and u is the term for errors. The unrestricted model takes into account all of the variables, without any concern for their significance as a regressor of US GDP. Although the adjusted R-squared value of this regression is quite high, .9998, this could be slightly inflated by insignificant variables. In addition, the coefficients of the significant variables could be greatly affected by the inclusion of these variables. Because of this, after completing the regression, it is important to test all of the variables by using either the p-value hypothesis test or t-statistic hypothesis test to see if the variables are significant at a 5% level, indicating confidence of 95%.

After completing these tests, two of the variables, US unemployment, and time are found to be insignificant. What this means is that within the regression model above, time and US unemployment do not consistently and accurately predict changes in U.S. GDP. In more detailed terms, this means that GDP changes over time were not consistently linear, and varied greatly from one year to the next, making time insignificant due to a high level of error. In a way, this is to be expected in a volatile period such as the recession since GDP changed sharply from year to year, and even fell in 2008 and 2009. US unemployment being insignificant means that increases in the percentage of unemployed workers in the US are unable to accurately predict changes in US GDP. This is an extremely interesting conclusion as the unemployment is a number that is constantly talked about as being a barometer of the health of the overall economy. It is important to know that these variables may be significant within other regression models or

even on their own, but when included as variables in the unrestricted model above, they are not significant, and as such, can be removed from the model.

Upon removing both time and US unemployment from the regression model, we are left with the restricted model as follows:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_3rUSGDP + \hat{\beta}_4lagUSGDP + u$$

This restricted model has simply removed the two insignificant variables from the regression in order to create a more accurate model. The statistics for this regression (as shown in Table 1 on page 3) show that the adjusted R-squared value has decreased very slightly to .9997, yet this model is more accurate as both variables, US GDP growth rate and US lagged GDP, are significant. The value of this year's GDP is explained by prior year's GDP and the growth rate of GDP over the previous ten years. This equation becomes a benchmark equation of sorts for US GDP. The goal of the other regression models is to see how similar variables in other nation's predict US GDP. This prediction is a rough way to show the degree of interdependence between foreign economies and the US economy based upon the previous ten years, including the recessional period.

By looking at the coefficients of this model, the more volatile of the two values is the GDP growth rate, in that each percentage change in the growth rate results in a massive change in the GDP for the next year, again something that we would expect. As growth rates are trending upward, we would expect this to have a consistent effect over time on the next period's GDP values, and consequently the same in periods of negative GDP growth, like the 2008 recession.

Finally, the coefficient for US lagged GDP is slightly greater than one, indicating that this model predicts a slight constant increase in GDP over the prior period. This coefficient is

signifying that over time, US GDP can be predicted fairly accurately simply by using the prior period's GDP multiplied by some factor. Again, this factor is fairly consistent, unlike US GDP growth rate which is expected to be more volatile in predicting US GDP.

All in all, what this final restricted model shows us is that the historical domestic economic data is a good predictor of changes or trends in US GDP. The adjusted R-squared value for the restricted model tells us that this regression equation is able to explain more than 99% of the changes in US GDP. While this model does not tell us exactly as to what is affecting the final US GDP, the variables making up the final equation are able to accurately and consistently predict changes in GDP. While this model may not provide accurate forecasts for future US GDP, it is a very good ex-post tool to use to demonstrate how changes of individual variables could have an effect on the domestic economy as a whole. Similarly, the following models will show the level of interdependence between the US and foreign economies. The conclusions will not be forecasting models, but rough snapshots that explain which countries have the strongest ties to the US economy.

Having looked at the important domestic values that have an effect upon US GDP and formulating a model based on these, the following regression models look at variables from the foreign countries that are being investigated to see their prediction abilities of US GDP. I will focus on the three most relevant countries' regressions, India, Germany, and Mexico, and evaluate the significant variables involved within them.

India Regression Model

Currently, India represents one of the largest developing economies in the world with a growing population and increase in foreign investment. Second to China, India is the next largest Asian country as well as second in terms of economic growth. The regression model for India examines three variables of the Indian economy: India GDP, India unemployment rate, and India GDP growth rate and how they can predict US GDP values. The adjusted R-squared values for these regressions were the third highest of all the countries investigated, and the results of the regressions can be seen below:

India Regression Results, Table 2:

Variable	Unrestricted Model	Restricted Model
India GDP	3.695* (0.37)	3.967* (0.38)
India Unemployment	-2.96E+11* (1.02E+11)	-2.89E+11* (1.13E+11)
India GDP Growth	-4.04E+09 (1.77E+10)	N/A
Constant	1.18E+13 (8.94E+11)	1.13E+13 (9.83E+11)
Adjusted R-squared	0.926	0.922

n=10, numbers in parentheses are standard errors

** indicates significance at 5% level*

Data for group can be found in Appendix A

These regressions look at the period from 2001-2011 with changes in US GDP as the dependent variable. The unrestricted model looks at three of the variables: the value of India’s GDP, the percentage of unemployment in India, and the percentage growth rate of India’s GDP from one period to the next. The equation for the unrestricted model above looks like this:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1INGDP + \hat{\beta}_2INunemployment + \hat{\beta}_3rINGDP + u$$

The adjusted R-squared value for this model is .926, indicating that changes in these variables are able to account for more than 92% of the changes in US GDP. This indicates a fairly high level of correlation must be present between the Indian economy and the US economy for these factors to have such high explanatory power of US GDP although not as strong as the US on US specification. As with all unrestricted models, it is important to verify if all of the variables included in the model are truly significant in order to be as accurate as possible. Again, by using a p-value test on the probability of the variables at a 5% significance level, we are able to restrict our model further.

In the case of the regression model for the Indian economy, we discover that the India GDP growth rate variable is in fact insignificant to this model. In this model, the variable of Indian GDP growth has an extremely high level of error, and as such is not able to consistently and accurately predict changes to US GDP. This most likely occurs within the recession years that were investigated because during these years, India's economy was still experiencing levels of GDP growth, where as the US economy had negative a negative growth rate. For example, in 2009 when the US GDP was declining, India's GDP had a growth rate of over 10%, a huge contrast. Similarly, since India's economy has grown at an extremely high rate over the last decade and has been somewhat volatile in the magnitude of its growth rate, it would not make a very good regression variable for the level of US GDP. Again, India's GDP Growth rate may pertain to other possible regressions, but when attempting to predict simply the level of US GDP, it is an insignificant variable and should be removed from the model.

After removing India GDP growth rate from our original regression model, we are left with the following final restricted regression equation:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 INGDP + \hat{\beta}_2 INunemployment + u$$

The restricted regression model above uses the two variables, India GDP and India unemployment rate to predict the levels of US GDP. Although the adjusted R-squared for the restricted model has fallen slightly to .922 (See Table on Page 7), our included variables have become more consistent as they are both significant estimators of US GDP.

By analyzing the coefficients of this final model, we see some fairly standard results based upon the order and magnitude of each variable in the model. First, by looking at the constant we can see that the regression model is expected to have a positive value, independent of the values of the other variables in the model. This is to be expected and shows that the relationship between the estimators of India's economy and US GDP has a fairly stable base level that is then affected by changes in the other estimators.

The second part of the model, India GDP has an estimated coefficient of just less than four. This coefficient shows that US GDP has fairly been greater than India's level of GDP by a certain multiplier. While this number has changed over the years as India's GDP has increased greatly due to their growing economy, the factor of four provides a consistent estimator as to what US GDP is in relation. Based on this variable's coefficient, we can expect that over time, as India GDP increases, US GDP would also increase at a fairly stable rate.

The final variable in this regression equation is the percentage of unemployment in India. As would be predicted, the estimated coefficient for this variable is negative, meaning that each percentage increase in the unemployment rate in India would decrease the expected level of US GDP. While it would not be expected that changes in India's unemployment rate would directly have an effect on US GDP, it is highly plausible that significant changes in India's unemployment rate would be mirrored in the US. Thus, meaning that any large change in this variable would be a very broad indicator of either increases or decreases to the strength of the

world economy. Since the US and India share high economic correlation, we would expect for similar changes to be experienced

What we can learn from this final, restricted regression model of India is that Indian economic factors are a good estimator of the strength of the US economy. In the final restricted regression model, changes in the factors of India's economy that were investigated are able to explain more than 92% of the changes in US GDP. Again, this model is not saying that India is directly responsible for changes in the US economy, but rather the two share an extremely high correlation, which is a sign of the increasingly global economy.

What this model lacks though, is any great way of accounting for an economic downturn, since India's economy has held a positive growth rate consistently for the past decade. Unless India were to experience a rapid, large increase in the unemployment rate, (which would predict a similar increase to the US unemployment rate) this model would never predict for the level of US GDP to decrease, such as what happened during the 2008 recession. As such, it is difficult to use a developing country such as India to predict the economy of a global power, such as the US. Overall, this model does show the increasingly global nature of the economy between two countries that share such a great geographical distance, but lacks the ability to account for economic downturns. The next regression looks at another global economic power, Germany, and its ability to predict changes in US GDP.

Germany Regression Model

Throughout the second half of the 20th century, Germany has been seen as an example of economic stability and efficiency. Even with the economic troubles that the European Union has experienced, Germany has been the economic leader and the stalwart country with a high level of economic prosperity. However, Germany did experience an economic downturn during the recession years of 2008 and 2009, and have been slow to recover, similar to the US. For the regression model of Germany, three economic metrics are used: Germany GDP, Germany unemployment, and Germany GDP Growth to see their ability to predict changes in US GDP. Similar to the previous models, an initial unrestricted model was used and then modified based upon insignificant variables. The results of the two regression models and their estimates can be seen below:

Germany Regression Results, Table 3:

Variable	Unrestricted Model	Restricted Model
Germany GDP	2.51* (0.37)	2.728* (0.23)
Germany Unemployment	-1.03E+10 (1.33E+11)	N/A
Germany GDP Growth	-3.02E+10 (2.03E+10)	N/A
Constant	5.81E+12 (1.94E+12)	4.95E+12 (6.88E+11)
Adjusted R-squared	0.8978	0.9311

n=10, numbers in parentheses are standard errors

** indicates significance at 5% level*

Data for group can be found in Appendix A

The unrestricted model uses all three variables and shows their ability to predict changes in US GDP. Germany GDP is the value of GDP at a given time in Germany, Germany unemployment is the percentage rate of unemployed workers in Germany at a given time, and

Germany GDP growth is the growth rate of GDP in Germany from one time period to the next.

The equation for this regression model is as follows:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 GERGDP + \hat{\beta}_2 GERunemployment + \hat{\beta}_3 rGERGDP + u$$

The adjusted R-squared value for this regression model is .8798, meaning that based upon changes to the independent variables, this model is able to explain almost 88% of the changes in US GDP. Again, this is a very high level of correlation between the German economy and US GDP due to the amount of explanatory power of the independent variables. However, upon examination, there are multiple variables included in this model that are not significant, and provide no explanatory power to US GDP. Because of this, we must perform statistical tests in order to see if they can be removed from our initial model.

After the tests, it is verified that two variables, Germany unemployment and Germany GDP growth rate, are both insignificant due to their high level of error. First, the estimated coefficient for Germany unemployment is negative as we might expect for increases in the German unemployment rate to mirror negative changes in the US economy, but it has a large amount of error. Since the standard error for this term is much larger than the estimated coefficient itself, this variable cannot be considered a significant predicting factor of US GDP. This is most likely due to unemployment being fairly constant, despite changes in US GDP over the similar period. In fact, Germany's unemployment rate has decreased consistently since 2006, while GDP levels have experienced high amounts of volatility over the same period, creating a high amount of error.

In the case of Germany GDP growth rate, the estimated coefficient is negative, signifying that increases in the growth rate would signal a decline in the estimate for US GDP. This is opposite of what we would predict as the growth rate of the German economy would seem to be

fairly correlated to increases in the level of US GDP. Again, this variable has a high amount of standard error associated with it, and as such, is not a consistent estimator of US GDP. The reason for this variable being insignificant is that based on the data collected, Germany GDP growth rate has been much more volatile over the years than US GDP, creating error in the prediction values.

Since these two variables are not significant in estimating the level of US GDP, we can remove them from our regression model and we are then left with this restricted regression equation:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 GERGDP + u$$

This equation is a simple linear regression model with Germany GDP as the only independent variable used to predict changes in US GDP. As opposed to the previous restricted regression models, this model improves upon the adjusted R-squared value by removing the insignificant variables. The adjusted R-squared value for this model is .9311, showing a very high level of correlation and explanatory power between Germany GDP and US GDP. This would be expected as Germany and the US are often thought of as the two strongest and most stable developed economies. By removing the variables of Germany unemployment and Germany GDP growth rate, this model removes a large amount of error associated with the equation, making it more accurate and leading to the much higher adjusted R-squared value.

By looking at the estimated coefficient for this model, the first term, the constant, has a high order of magnitude which we would expect due to the extremely high values of both countries' GDP. This indicates that no matter what the value of our independent variable, Germany GDP, the level of US GDP should remain a consistently high value.

The estimated coefficient of Germany GDP is a much smaller value, 2.7, that acts as a factor or multiplier for the difference in magnitude between Germany GDP and US GDP. Although this number may not be accurate every year, the low amount of standard error signifies that it is a very good estimate of US GDP in any given year. This variable is very significant, as expected, and provides a high amount of explanatory power for US GDP due to the high correlation between the economies of the two countries.

Both the US and Germany have extremely strong market based economies that have been fairly stable over the course of the late 20th and early 21st century. Both are easily the strongest economies within their continent and region, and as such have the strongest financial markets. Since the two economies have so much in common, it is no surprise that they should be so highly correlated in both periods of growth and recession. While other variables most certainly have a factor in determining the level of both Germany and US GDP, both would be fairly accurate predictors and indicators of changes in the other's economy. As the global economy becomes even more connected, the US can use Germany GDP as a predictor of impending changes that may be experienced in the market, such as the US experiencing the effects of the European debt crisis. While the internal US regression has yielded a better prediction ability as expected, the Germany model shows an example of extremely high correlation to a developed foreign economy. The next regression that will be examined in this project shows the effects of close regional proximity and trade, their ability to predict changes to US GDP.

Mexico Regression Model

Whenever countries share borders, it is likely that their economies share many similarities. Often trade and production are shared between the countries and as such, their economies become extremely correlated. Since the creation of NAFTA, this has been true of the relationship between Mexico and the US. With the abolishment of tariffs, there has been a large increase in economic activity between the two countries. This influx of trade and movement of capital across the border links the two countries' economies further as the economic climate in one country will normally mirror changes in the other. Throughout the recession years, this was extremely noticeable as the economic shock was widespread between the two countries. As such, these models below focus on the economic metrics in Mexico and their ability to explain changes in US GDP. Three regression models were used in this equation and their results can be found below:

Mexico Regression Results, Table 4:

Variable	Model 1	Model 2	Model 3
Mexico GDP	7.893* (.427)	7.605* (0.464)	8.834* (0.641)
Mexico Unemployment	3.80E+11* (8.13E+10)	4.76E+11* (8.55E+10)	N/A
Mexico GDP Growth	-1.46E+10* (6.23E+9)	N/A	-2.85E+10* (1.07E+10)
Constant	4.55E+12 (3.60E+11)	4.36E+12 (3.62E+11)	5.22E+12 (5.84E+11)
Adjusted R-squared	0.9865	0.9834	0.9544

n=10, numbers in parentheses are standard errors

* indicates significance at 5% level

Data for group can be found in Appendix A

The unrestricted regression for Mexico uses all of the above variables to see their ability to explain changes in US GDP. The three variables used are Mexico GDP which is the value of

Mexico's GDP at a given time, Mexico unemployment which is the percentage number of unemployed workers in Mexico at a given time, and Mexico GDP Growth which is the percentage value of growth in Mexico's GDP from one period to the next. The equation for this first model looks like this:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 MEXGDP + \hat{\beta}_2 MEXunemployment + \hat{\beta}_3 rMEXGDP + u$$

The adjusted R-squared value for this first regression model is .9865, signifying that changes in the independent variables are able to explain 98% of changes in US GDP. This extremely high value represents that these variables for Mexico's economy have a large amount of explanatory power for US GDP. This is most likely due to the high level of trade and cross-border interaction of the two countries considering their extremely close proximity. Also, unlike most of the previous models, Mexico's model has no insignificant variables included. Thus, rather than removing the insignificant variables and looking at the restricted model, two additional models were reviewed by removing one variable from each and examining changes to the estimated coefficients.

First, in order to have a tool for comparison, it is necessary to evaluate the estimated coefficients from the first regression model. The first variable, Mexico GDP has an estimated coefficient of 7.89, meaning that any increases in Mexico's GDP are predicted to coincide with changes to US GDP at the factor of 7.89. The coefficient of the constant, or intercept term is positive and large, indicating that no matter what changes occur in the independent variables, the US level of GDP is expected to remain large and fairly constant as well.

The final two coefficients are not what we would expect for them to be, and can then cause some problems. First, Mexico unemployment has a coefficient that is positive, meaning that increases in Mexico's unemployment rate will predict increases to the US level of GDP.

This would be expected to be negative, but since Mexico’s unemployment rate has increased fairly constantly over time, it yielded a result opposite of what we would predict. Similarly, Mexico GDP Growth is negative, meaning that increases in the growth rate will predict decreases to US GDP. This again is not expected, but could be due to the higher level of volatility with regards to Mexico’s GDP Growth. These are two theories, but these values do provoke thought as to the economic relationship between the two countries. To examine these variables further, two additional models have been created where one variable is removed in each to see if there is any drastic change in the estimated equation.

The first of the two models sees the variable of Mexico GDP Growth removed from the equation, making the new regression equation the following:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 MEXGDP + \hat{\beta}_2 MEXunemployment + u$$

The adjusted R-squared value for this regression equation is mostly unchanged at .9834, dropping just slightly. Since the adjusted R-squared value for this equation is not very different from the original model, it shows that the omitted variable, Mexico GDP Growth, is most likely the least significant variable of those in model.

This can also be affirmed by looking at the changes of the other independent variables compared to the initial model. As seen in Table 4, the constant term was largely unchanged, with only a slight decrease in its estimate, but no major change. Similarly, the estimated coefficients of Mexico GDP and Mexico unemployment decreased and increased respectively very slightly, along with their error terms remaining close to their original values. Since these values, along with the adjusted R-squared value for the regression, remained largely unchanged, this signifies that the model is consistent and these two variables have the most explanatory

power. Another way to confirm this is by returning to our original regression model and removing a different variable and examining the new model's results.

For the third and final model regarding Mexico's economy, Mexico unemployment is removed from the equation, leaving the following regression model:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 MEXGDP + \hat{\beta}_2 rMEXGDP + u$$

For this final model, the adjusted R-squared value has decreased noticeably from the original model, down to .9544, as seen in Table 4. This decrease from the unrestricted model shows that the removed variable has a large amount of explanatory power, causing the regression model to be less accurate. Since this model is not as accurate at predicting US GDP, the variable that was removed, Mexico unemployment must be significant for this regression model. A final way to affirm that the variable is extremely significant is looking at the estimated coefficients in the regression results and identifying any large changes or increases in error.

Upon reviewing the coefficients, all of the variables experienced changes that are quite different from the original model. First, the constant term for the third model increased a decent amount, although not an increase in magnitude, but the error term for it increased quite noticeably. The combination of these two increases can be seen as a significant change and one that needs to be considered. Next, the most important independent variable for the equation, Mexico GDP, experienced increases in both its coefficient and error. While neither term increased by an order of magnitude, both are different from the first two models where the terms remained fairly consistent. Finally, the final independent variable in this regression model, Mexico GDP Growth, increased the absolute value of its estimated coefficient and the amount of error which increased by an order of magnitude. Since all of the variables experienced some significant change, it can be determined that this model is not nearly as accurate as the previous

models and should not be seriously considered. By looking at all three of the models, the best regression model for predicting and explaining changes in US GDP is the first unrestricted model containing all the variables considered with Mexico's economy.

The two countries' economies have grown to become highly reliant upon one another over the past decade, and this is reflected in the final regression model. Both countries experienced dips in economic output during the recessional years and have recovered at slightly similar rates as well. It will be interesting to see in the future if both countries continue to overcome the recession, or if one or both succumb to periods of poor economic health once again. The following regression model is an example of a poor predictor for US GDP to provide a contrast to the previous models.

China Regression Model

Of the countries researched, the poorest predictor of US GDP based on the regression models was China. While this may come as a shock due to China’s rapidly increasing economic prosperity, upon further investigation it makes some amount of sense. China has experienced large amounts of volatility over the past decade, and as such probably was not extremely correlated to the US economy. The unrestricted model for China looks at three economic metrics and their ability to predict changes in US GDP. This model was then restricted to only the significant variables. The regression results can be seen in the table below:

China Regression Results, Table 5:

Variable	Unrestricted Model	Restricted Model
China GDP	.7645* (0.207)	.7532* (0.119)
China Unemployment	-3.99E+11 (5.35E+11)	N/A
China GDP Growth	7.93E+11 (6.34E+11)	N/A
Constant	1.08E+13 (2.26E+12)	1.03E+13 (4.67E+11)
Adjusted R-squared	0.768	0.795

n=10, numbers in parentheses are standard errors

* indicates significance at 5% level

Data can be found in Appendix A

The first model, which includes all three of the variables, has an adjusted R-squared value of only .768, meaning that changes in the independent variables can only explain 76% of changes in US GDP. The variables included in this model are China GDP, the value of China’s GDP at a point in time, China unemployment, the percentage of unemployed workers in China at a point in time, and China GDP Growth, the percentage growth rate in Chinese GDP from one period to the next. The equation for the model looks like this:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 CHINAGDP + \hat{\beta}_2 CHINAunemployment + \hat{\beta}_3 rCHINAGDP + u$$

This first model for China is not a very good predictor of US GDP and has two variables that are insignificant predictors of US GDP. The likely reason for this model not providing much explanatory power for US GDP is the high amount of volatility and rapid growth of China. Compared to the US, China's economy has been constantly changing and growing whereas the US economy, save the recession, was fairly stable. In order to see the proper final regression for China, the model must not include the two insignificant variables, China unemployment and China GDP Growth rate. After removing these two variables from the model, the equation looks like this:

$$USGDP = \hat{\beta}_0 + \hat{\beta}_1 CHINAGDP + u$$

The adjusted R-squared value for this restricted regression model is .795, showing improvement over the unrestricted model, but still not a very accurate predictor of US GDP. While for this experiment, China is not a good predictor of changes to US GDP based upon the metrics reviewed, in the future this could change. As China's growth begins to level off and become more constant, the two economies will be more highly correlated, changing China's regression model and prediction power. The same goes for the other countries that yielded poor regression models. While over the time period surveyed their economies did not show high correlation to the US, all of this could change in coming years.

Conclusion

Having compiled data and created regression models for many different possible variables, the four discussed here previously have stood out as being the most viable models to explain changes in US GDP. These models are derived mostly by using variables related to foreign economies, and one that uses domestic data to create a baseline model. The models that were discussed here have proven that like the US, many countries around the world have experienced similar economic downturns over the past years, and that the global economy is becoming increasingly interconnected. As such, we will have to be aware of possible warning signs of economic problems in the future and find ways to forecast and prepare for these appropriately.

While these models are not meant to be used solely for forecasting purposes as they are very general models, they are able to show which countries the US shares high levels of economic correlation with. Similarly, these models will allow us to use different variables to help predict the sign of changes in the US economy, i.e. increases or decreases in GDP.

It must also be considered that these are only the best models from the selected data. There may be other countries that apply to this study, but were not considered as a part of this research. Similarly, there are many other countries that were researched whose results were abysmal in terms of creating a viable model in regards to US GDP. Furthermore, all the data and results of research can be found in the attached Appendix A.

As the economy continues to grow and change, it becomes increasingly important for policymakers and investors alike to use econometric principles to forecast in a similar manner to this. Finding high levels of correlation can lead to important revelations and allow for more accurate and pertinent policies to be made in the future. While they may be unable to prevent the

economy from falling into a recession, the increased knowledge will allow for quicker reactions and better preparation in the long term. We may never know the future, but the closer we are with our predictions, the more our present actions can set us up for success.

Works Cited

"Data." *The World Bank*. The World Bank Group, 01 Apr 2013. Web. 1 April 2013.

<<http://data.worldbank.org/>>.

"IMF Data and Statistics." *International Monetary Fund*. International Monetary Fund, n.d.

Web. 2 May 2013. <<http://www.imf.org/external/data.htm>>.

Appendix

Attached to this project is a flash drive containing the important data used in creating this report. There are many different files, ranging from the raw data files obtained from World Bank and IMF, as well as graphs, charts, and other measurements that were not used directly in this report. All of the data used can be found in the file labeled “Sorted Data” where the raw data of all ten countries is sorted into tables and graphs based upon the different metrics. Finally, all of the regression data can be found on the Eviews work file named “Regressions”. In order to view this data, you must have a copy of Eviews software, or have access to Ball State’s network. Again, the Eviews file provides more data than what is discussed directly in this report. In addition, the tables shown in the report have been re-formatted and will not resemble the tables and information that will be found in the work file.